Regulation of HIV-1 splicing
Müller, Nancy

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
The HIV-1 Tat protein enhances splicing at the major splice donor site

Nancy Mueller, Bep Klaver, Ben Berkhout, Atze T. Das

Laboratory of Experimental Virology, Department of Medical Microbiology, Center for Infection and Immunity Amsterdam (CINIMA), Academic Medical Center, University of Amsterdam, The Netherlands

Manuscript in preparation
Abstract
Transcription of the HIV-1 proviral DNA and subsequent processing of the primary transcript results in the production of a large set of unspliced and differentially spliced viral RNAs. The major splice donor site (5’ss) that is located in the untranslated leader of the HIV-1 transcript is used for the production of all spliced RNAs and splicing at this site has to be tightly regulated to allow the balanced production of all viral RNAs and proteins. We demonstrate that the viral Tat protein, which is known to activate viral transcription, also stimulates splicing at the major 5’ss. Like for the transcription effect, Tat requires the viral LTR promoter and the TAR RNA hairpin for splicing regulation. These results indicate that HIV-1 transcription and splicing are tightly coupled processes through the coordinated action of the essential Tat protein.
Introduction
The retrovirus HIV-1 has a ~9.7 kb RNA genome that upon infection of cells is reverse transcribed into double-stranded DNA. This DNA integrates into the cellular genome and can subsequently be transcribed by RNA polymerase II (RNAP II), which results in the production of new viral RNA transcripts. The full-length 5’-capped and 3’-polyadenylated transcript is used as messenger RNA (mRNA) for translation of the Gag and Pol proteins and as genomic RNA that is packaged into new virus particles. The primary transcript contains several 5’ and 3’ splice sites (5’ss and 3’ss) and differential usage of these sites during the co-transcriptional splicing process results in a large variety of spliced mRNAs that are translated into the other viral proteins, including the Tat protein that stimulates HIV-1 transcription (1).

The RNAP II transcription complex that initiates at the 5’LTR promoter in the absence of Tat is paused at the stable TAR (transacting responsive) RNA structure formed at the 5’ end of the nascent transcripts, which causes premature termination of transcription (2, 3). As a consequence, predominantly short transcripts are produced (4). Binding of Tat to TAR results in the recruitment and activation of the positive transcription elongation factor B (P-TEFb). P-TEFb preassembles at the U3 region of the LTR promoter in a catalytically inactive state bound to the inhibitory 7SK small nuclear ribonucleoprotein (snRNP) (5-7). Tat mediates transfer of P-TEFb to TAR RNA, resulting in displacement of 7SK snRNA. The cyclin-dependent kinase 9 (CDK9) component of P-TEFb can subsequently activate RNAP II through phosphorylation of the C-terminal domain (8-10). P-TEFb also phosphorylates the negative elongation factors (NELFs), which causes their release from the transcription complex, and P-TEFb directs the recruitment of TATA-box binding protein to the LTR promoter, which stimulates the assembly of new transcription complexes (11). Tat may also recruit several chromatin-modifying proteins to remodel the promoter region (reviewed in (9, 12, 13)). In addition, Tat has been reported to stimulate phosphorylation of several transcription factors that bind the LTR promoter, like Sp1 and NFκB (reviewed in (12)). These combined Tat effects result in the formation of a more processive transcription complex that is not paused at the TAR RNA structure. Thus, Tat activates the production of the complete set of unspliced and spliced viral RNAs needed for virus replication.

HIV-1 splicing has to be tightly regulated for the balanced production of all viral RNAs and proteins. The major 5’ss in the 5’ untranslated leader is used for the generation of all spliced mRNAs. The level of splicing at this 5’ss is controlled by several factors. First, splicing regulatory elements that bind splicing enhancing or silencing factors are located in close proximity to the 5’ss (14, 15). Second, incomplete sequence complementarity between the 5’ss nucleotides (nts) and the U1 snRNA results in suboptimal spliceosome assembly during the first step of the splicing process (16). Third, the sequence surrounding the major 5’ss can fold a repressive hairpin structure. We demonstrated that destabilization of this splice donor (SD) hairpin increased splicing, whereas stabilization reduced splicing, which indicates that splicing is restricted by the RNA structure, most likely by limiting the access of splicing factors, including U1 snRNA and regulatory factors (17, 18).
Transcription and splicing are tightly coupled processes (19). Several studies suggested a role for transcription factors in HIV-1 splicing and a role for splicing factors in HIV-1 transcription. The transcription elongation regulator 1 (TCERG1, also known as CA150) that associates with the RNAP II complex and is involved in Tat-activated HIV-1 transcription (20) also binds to the splicing factor SF1 (21). SF1 was originally identified as a constitutive splicing factor (22, 23), but other studies indicated that it represses transcription (24, 25). The Tat stimulatory factor Tat-SF1 was first discovered as a cellular protein required for activation of HIV-1 transcription (26-30). Several studies indicated that Tat-SF1 associates with RNAP II and other factors including Tat, P-TEFb and TCERG1 (CA150) (26, 31-35). More recent studies, however, demonstrated that Tat-SF1 is not required for Tat activation of transcription, but rather influences the ratio of unspliced to fully spliced HIV-1 RNAs (36, 37). Tat-SF1 interacts with several components of the spliceosome, including U1 snRNP proteins and all five splicosomal U snRNAs (38). The human splicing factor SKIP (splicing-associated c-Ski-interacting protein) also influences both HIV-1 transcription and splicing. SKIP associates with P-TEFb and subsequently stimulates transcription elongation. SKIP also associates with U5 snRNP proteins and tri-snRNP110K, but not with other splicing factors (including Tat-SF1 and TCERG1/CA150), and enhances utilization of an alternative 3'ss for a Tat-encoding RNA (39). Recently, the splicing factor SRSF1 (SF2/ASF) and Tat were shown to recognize overlapping sequences within TAR and 7SK RNA (40). A model was proposed in which the ubiquitously expressed SRSF1 can activate basal HIV-1 transcription in the absence of Tat by recruiting P-TEFb to TAR from the 7SK snRNP. When present, Tat can substitute for SRSF1, which results in more efficient transcription elongation.

The Tat protein has a major role in the regulation of HIV-1 transcription, but other functions have been suggested to support HIV-1 replication. For example, Tat may influence reverse transcription, capping and translation of the viral RNAs (41-45). Tat also influences 3'ss selection and promotes usage of the sites for the Rev and Env/Nef specific mRNAs (41). We here investigate the role of Tat in splicing at the major SD site that is used for the production of all spliced RNAs.

Results

Tat activates splicing at the major 5'ss

To analyze the effect of Tat on splicing at the major 5'ss site, we used a simple reporter construct in which the HIV-1 5’ LTR-promoter and leader RNA sequence, including the major 5'ss (SD<sup>wt</sup>: CUG▲GUGAGUAC; ▲ indicating the splice site), was fused to the firefly luciferase gene that encodes a cryptic 3'ss (Fig. 1A; (17)). Upon transfection of this LTR-leader-luc construct in cells, transcription and subsequent RNA processing will result in the production of unspliced RNAs that encode luciferase and spliced RNAs that do not. These RNA products can be detected by reverse transcription followed by PCR amplification (RT-PCR), which results in 710-bp and 429-bp products for the unspliced and spliced RNAs, respectively (Fig. 1A). Moreover, the intracellular luciferase level can be measured to quantify splicing. We previously used this reporter to determine the effect of local RNA structure and binding of
splice regulatory SR proteins in the SD region on splicing frequency (14, 18).

The LTR-leader-luc plasmid was transfected into 293T cells together with either the Tat-expressing pTat plasmid or the empty pcDNA3 vector. Transcription from the LTR promoter is activated by the Tat protein and we indeed measured an increased luciferase level with pTat (14, 18) (Fig. 1B). Both in the absence and presence of Tat, unspliced and spliced RNAs were detected by RT-PCR, but the ratio between these products varied. The unspliced RNA product was more abundant in the absence of Tat, whereas the spliced RNA product was more abundant in the presence of Tat (Fig. 1C). A similar pattern was observed when the cells were transfected with smaller amounts of the reporter construct (data not shown) or when the PCR analysis was repeated with different dilutions of the cDNA input (Fig. 1C). These results indicate that the observed splicing effects are not due to quantitative differences in cDNA input resulting from Tat-induced activation of transcription, but likely reflect Tat activation of splicing at the major 5'ss site.

Figure 1. Tat activates splicing at the HIV-1 major 5'ss. (A) The HIV-1 provirus (DNA) is shown with its open reading frames and long terminal repeat regions (LTRs). The 5' and 3' LTR consist of U3, R and U5 regions. Transcription starts at the U3-R border (+1) of the 5' LTR and the transcripts are polyadenylated at the R-U5 border of the 3' LTR. The major 5'ss is used for the production of all spliced RNAs. In the pLTR-leader-luc construct, the HIV-1 5' LTR and leader region (including the major 5'ss) are coupled to the firefly luciferase gene that encodes a cryptic 3'ss. The primers used for RT-PCR analysis of the RNA products are indicated. This RT-PCR will result in a product of 710 bp for the unspliced RNA and 429 bp for the spliced RNA. (B) 293T cells were transfected with the pLTR-leader-luc constructs with or without a Tat-expressing plasmid. The intracellular luciferase production was measured after 48 h. The mean, standard error and fold-induction of 3 independent measurements are shown. (C) The unspliced and spliced RNAs were analyzed by RT-PCR. Different dilutions of the reverse transcription cDNA product were used as input in the PCR (1x, undiluted). The DNA size marker is shown at the left side (bp).

The Tat splicing effect does not depend on local RNA structure or SR protein binding

The sequence surrounding the major 5'ss can fold an RNA stem-loop structure, the SD hairpin, that exposes the splice site in the loop (Fig. 2A). To analyze the role of this hairpin structure in Tat activation of splicing, we analyzed the RNAs produced by LTR-leader-luc constructs with different nucleotide (nt) substitutions that destabilized the RNA structure but did not affect the 5'ss sequence (Fig. 2A). These mutations, except for the A mutation, also removed some of the putative SR protein binding sites, in particular putative sites for the splicing inhibitory SRSF2 (SC35) protein (Fig. 2B). As previously described, most of these mutations slightly increase the splicing efficiency.
- both in the absence and presence of Tat - and resulted in a reduced luciferase level when compared with the wild-type (wt) construct (14, 18). RT-PCR analysis of the RNAs produced by the different variants revealed that the unspliced RNA product was nevertheless more abundant in the absence of Tat and the spliced RNA product was more abundant in the presence of Tat (Fig. 2C, upper and lower panels, respectively). In fact, the Tat-induced increase in splicing efficiency was more pronounced than the mutation-induced increase. These results demonstrate that destabilization of the SD hairpin or removal of the putative SR protein binding sites did not impede Tat activation of splicing. In other words, the Tat splicing effect does not depend on local 5’ss characteristics like hairpin structure and SR protein binding.

Disentanglement of the Tat transcription and splicing effects
These results suggest that Tat not only activates transcription from the LTR promoter, but also stimulates splicing at the major 5’ss. These Tat effects have opposing effects on the read out in LTR-leader-luc transfection experiments. Whereas an increase in transcription will boost the luciferase level, an increase in splicing will reduce this level. The observed increase in luciferase production observed upon Tat administration (Fig. 1B) thus reflects the sum of these opposing effects. To disentangle these transcription and splicing effects, we designed an LTR-leader-luc construct with an inactivated major 5’ss sequence (SD\textsuperscript{m}: CUG\textbulletGUGCGCAC, mutations underlined). The introduced SD\textsuperscript{m} mutation reduces the sequence complementarity between the 5’ss and U1 snRNA and - as previously demonstrated - completely blocks splicing at the major 5’ss (18). As a consequence, an increase in luciferase production measured upon Tat administration will solely reflect Tat-mediated activation of transcription. It has previously been described that binding of U1 snRNA to the 5’ss prevents polyadenylation at the polyadenylation signal present in the 5’ LTR (46, 47). We anticipated that the SD\textsuperscript{m} mutation could thus cause premature polyadenylation of the transcripts, which would result in very short RNAs (~97 nt plus polyA tail) and complicate the analysis. We therefore included SD\textsuperscript{wt} and SD\textsuperscript{m} constructs with a mutation in the 5’ polyadenylation signal (pA\textsuperscript{wt}: AAUAAA; pA\textsuperscript{m}: AACAAA) that inactivates this site (Fig. 3A).

293T cells were transfected with the LTR-leader-luc constructs with pTat or the empty pCDNA3 vector, and the RNA products and luciferase level were analyzed after 48 h. Analysis of the RNAs produced by the wt construct (SD\textsuperscript{wt}/pA\textsuperscript{wt}) demonstrated the typical shift toward spliced RNA products upon Tat administration (Fig. 3B). For the SD\textsuperscript{m}/pA\textsuperscript{wt} construct, only the unspliced RNA was detected both in the absence and presence of Tat, which confirms that the 5’ss mutation did completely block splicing. Accordingly, the SD\textsuperscript{m} construct consistently produced significant higher luciferase levels than the wt construct (Fig. 3C). Moreover, Tat activated luciferase expression 4.8-fold for the wt construct, whereas expression from the SD\textsuperscript{m} construct was induced 9.0-fold (Fig. 3D). This difference in fold-induction can be attributed to the Tat-splicing effect, which tempers Tat-induced luciferase production by the SD\textsuperscript{wt} construct. Mutation of the polyadenylation signal did not influence the RNA patterns nor the luciferase levels (compare SD\textsuperscript{wt}/pA\textsuperscript{wt} with SD\textsuperscript{m}/pA\textsuperscript{m} and SD\textsuperscript{m}/pA\textsuperscript{wt} with SD\textsuperscript{m}/pA\textsuperscript{wt} in Fig. 3B and 3C), which indicates that this site was not used for
Figure 2. Role of SD hairpin stability and putative SR protein binding sites in Tat activation of splicing.

(A) The sequence of the SD region with the putative SR protein binding sites (54) and the SD stem-loop structure with the introduced mutations are shown (▲, cleavage site; grey circles indicate the 5’ss nts to which the U1 snRNA can bind). (B) The presence (+) or absence (-) of putative SR protein binding sites, the thermodynamic stability of the SD hairpin (ΔG in kcal/mol) and the splicing efficiency (as previously determined; (16, 18)) of the wt and mutant SD sequences are shown. (C) 293T cells were transfected with the wt or mutant pLTR-leader-luc constructs without (upper panel) or with (lower panel) pTat and the RNA products were analyzed by RT-PCR.
polyadenylation in the SD<sup>wt</sup> and SD<sup>m</sup> transcripts. As observed for the pA<sup>wt</sup> constructs, Tat administration induced luciferase expression from the splicing-competent pA<sup>m</sup> construct (SD<sup>wt</sup>/pA<sup>m</sup>) to a lesser extent (4.1 fold) than from the splicing-defective variant (SD<sup>m</sup>/pA<sup>m</sup>) (9.3 fold; Fig. 3D), which confirms the Tat-splicing effect.

We calculated the splicing frequency of the different constructs in the absence and presence of Tat. For every SD<sup>wt</sup>/SD<sup>m</sup> set, the luciferase level observed for the SD<sup>m</sup> variant corresponds to the total RNA production (100% unspliced RNA). The luciferase level observed in mock transfected cells (which is close to zero) reflects the absence of unspliced RNA (0% unspliced RNA). These values were used to interpolate the unspliced RNA level of the SD<sup>wt</sup> variants (luciferase level SD<sup>wt</sup> construct / luciferase level SD<sup>m</sup> construct x 100%; Fig. 3D). In the absence of Tat, both the SD<sup>wt</sup>/pA<sup>wt</sup> and SD<sup>wt</sup>/pA<sup>m</sup> constructs produced ~40% unspliced RNA (~60% spliced RNA). In the presence of Tat, both constructs produced only ~20% unspliced RNA (~80% spliced RNA) (Fig. 3D), which illustrates the stimulatory effect of Tat on splicing at the 5'ss site.

Figure 3. Tat activates transcription and splicing. (A) LTR-leader-luc variants were constructed with a wt (SD<sup>wt</sup>, CUG>GUGAGUAC) or inactivated major 5'ss (SD<sup>m</sup>, CUG>GUGCGCAC) and a wt (pA<sup>wt</sup>, AAUAAA) or inactivated 5' polyadenylation site (pA<sup>m</sup>, AACAAA; wt and mutated elements indicated with filled and open symbols, respectively). The anticipated effect of Tat on transcription and splicing is indicated (+, effect; -, no effect). (B) 293T cells were transfected with the LTR-leader-luc constructs without or with pTat. The RNA products were analyzed by RT-PCR. (C) Analysis of the luciferase production. The mean and standard error of 2 independent experiments are shown. The fold-induction in luciferase expression observed upon Tat administration is indicated. (D) The percentage of spliced RNA was calculated for every mutant. The luciferase level of the SD<sup>m</sup> mutant (100% unspliced) was used to interpolate the unspliced RNA level for the SD<sup>wt</sup> constructs (luciferase level SD<sup>wt</sup> construct / luciferase level SD<sup>m</sup> construct x 100%).
**Tat requires TAR for its splicing effect**

Tat binds the TAR hairpin that is present at the 5’ end of viral transcripts to activate transcription at the 5’LTR. To test whether splicing activation also depends on the Tat-TAR interaction, we tested an LTR-leader-luc construct with several nt substitutions in TAR that prevent Tat binding (TAR\textsuperscript{m}; Fig. 4A). These mutations block the Tat-TAR transcription activation mechanism (48, 49). Whereas the wt construct (TAR\textsuperscript{wt}) revealed the Tat-induced increase in spliced RNA in transfected cells (Fig. 4B), the TAR-mutated construct (TAR\textsuperscript{m}) demonstrated an identical RNA pattern in the absence and presence of Tat, which indicates that Tat did not activate splicing. These results demonstrate that Tat requires a wt TAR element for its splicing effect. The TAR\textsuperscript{m} construct produced a similar level of luciferase in the absence and presence of Tat (Fig. 4C), which is in agreement with the Tat-independent promoter activity of this construct.

![Figure 4. Tat requires TAR to activate splicing.](image)

**Figure 4. Tat requires TAR to activate splicing.** (A) LTR-leader-luc constructs with a wt (TAR\textsuperscript{wt}) or mutant (TAR\textsuperscript{m}) TAR hairpin. The TAR\textsuperscript{m} construct contains several nt substitutions in the TAR loop and bulge domains to prevent Tat binding. This construct contains two binding sites (tet operators; tetO) for the doxycyline (dox)-inducible transcriptional activator protein rtTA in the U3 promoter region, which allows the exogenous activation of transcription by the administration of dox and rtTA. (B-C) 293T cells were transfected with the TAR\textsuperscript{m} construct, an rtTA-expressing plasmid and either pTat or pcDNA3, and cultured with dox. Cells were transfected with the TAR\textsuperscript{wt} construct and either pTat or pcDNA3 in parallel. The RNA products (B) and luciferase production (C) were analyzed at 48 h after transfection as described in Fig. 3.
The U3 promoter region is required for the Tat splicing effect

Tat preassembles with P-TEFb at the U3 region of the LTR promoter and is subsequently transferred to the TAR hairpin at the 5’end of nascent viral transcripts (5). To test whether or not the splicing-stimulating activity of Tat also requires the viral U3 promoter region, we designed luciferase reporter constructs in which the production of TAR-containing transcripts is driven by the non-related dox-inducible pTet promoter (Fig. 5A). Variant constructs with a wt or mutant TAR and a wt or mutant 5’ss sequence were transfected into cells and the produced RNAs and luciferase level were analyzed. As expected, RT-PCR analysis demonstrated that the SD\textsuperscript{m} variants produced only the unspliced transcript, both in the absence and presence of Tat (Fig. 5B). The SD\textsuperscript{wt} constructs produced both the spliced and unspliced RNAs. The ratio between these products was similar for the TAR\textsuperscript{wt} and TAR\textsuperscript{m} variants. Moreover, this ratio was not influenced by Tat, which indicates that Tat did not increase splicing.

Apparently, Tat does not only require the wt TAR hairpin, but also the U3 promoter region to activate splicing at the 5’ss. Tat did not increase luciferase expression of any of these constructs (Fig. 5C), which is in agreement with the U3 promoter requirement for Tat-mediated activation of transcription. Thus, Tat requires the U3 promoter and TAR elements not only for activation of HIV-1 transcription, but also for activation of splicing at the major 5’ss, which may indicate that these processes are coupled.

Figure 5. The U3 promoter region is required for the Tat effect on splicing. (A) In the tetO-leader-luc construct, transcription is driven by the pTet promoter. This dox-inducible promoter consists of 7 binding sites (tet operators; tetO) for the dox-inducible transcriptional activator protein rtTA coupled to a TATA-box containing minimal promoter sequence. Variants were constructed with a wt (TAR\textsuperscript{wt}) or mutant (TAR\textsuperscript{m}) TAR element and a wt (SD\textsuperscript{wt}) or mutant (SD\textsuperscript{m}) SD sequence. (B-C) 293T cells were transfected with the tetO-leader-luc constructs, an rtTA-expressing plasmid and either pTat or pcDNA3, and cultured with dox. The RNA products and luciferase expression was analyzed as described in Fig. 3. The mean and SD of 3 experiments are shown.
**Tat activates HIV-1 splicing**

Analysis of the Tat effect on splicing efficiency in the complete HIV-1 genome context is complicated by pleiotropic effects that Tat mutations may have on viral gene expression. Tat may not only affect transcription and splicing at the major 5’ss, but also other processes, like 3’ss selection (41). The fact that the regulatory Tat and Rev proteins are produced from fully spliced transcripts further complicates the analysis. Rev stimulates nuclear export of the unspliced and singly spliced viral RNAs. Moreover, Rev has been reported to influence other viral replication processes, like translation and RNA encapsidation (50). Thus, changes in the splicing efficiency of the complete HIV-1 genome will have complicated direct and indirect effects on RNA production and processing.

We nevertheless set out to confirm the stimulatory effect of Tat on splicing at the major 5’ss in the complete HIV-1 genome context. To prevent the Tat and Rev regulatory loops during viral gene expression, we constructed a Tat and Rev deficient HIV-1 variant (HIV-Tat\textsuperscript{m}/Rev\textsuperscript{m} in Fig. 6A). In this virus, the production of Tat was blocked by substitution of the Gly\textsuperscript{15} and Glu\textsuperscript{17} codons in the Tat open reading frame with translation stop codons that will cause premature termination of translation (51). To prevent Rev production, the translation start and Gly\textsuperscript{6} codons in the Rev open reading frame were replaced by translation stop codons. Upon transfection into 293T cells, this construct did not produce detectable viral Gag-derived CA-p24 protein, while viral gene expression could be rescued by co-transfection of Tat and Rev expressing plasmids (data not shown). Upon co-transfection of Rev plasmid and an increasing amount of Tat plasmid, a gradual increase in the CA-p24 level was observed (SD\textsuperscript{wt} in Fig. 6B). As the Gag precursor protein is translated from the unspliced viral transcript, the Tat-induced increase in the CA-p24 level reflects the combined effects of activation of transcription and stimulation of splicing at the major 5’ss, which have an opposite effect on the full length RNA and CA-p24 level. To disentangle these effects, we excluded the Tat effect on splicing by mutational inactivation of the major 5’ss site in HIV-Tat\textsuperscript{m}/Rev\textsuperscript{m} (SD\textsuperscript{m} in Fig. 6A). Upon transfection of 293T cells with 1000 ng of the SD\textsuperscript{wt} and SD\textsuperscript{m} constructs, we observed reduced CA-p24 levels for the SD\textsuperscript{m} variant compared to the SD\textsuperscript{wt} virus (Fig. 6B). This reduction is most likely due to the induction of polyadenylation at the 5’ polyadenylation site, as U1 snRNA binding to the major 5’ss is required to prevent premature polyadenylation in the complete HIV-1 context (46, 47) and our 5’ss mutation was designed to reduce U1 snRNA binding. The short RNAs resulting from this premature polyadenylation do not encode Gag protein, which can explain the relatively low CA-p24 level.

Analysis of the viral transcripts by RT-PCR, which allowed the simultaneous detection of the unspliced RNA and the spliced Vif-encoding RNA (Fig. 6C), demonstrated that the SD\textsuperscript{wt} virus produced both unspliced and spliced RNAs, whereas the SD\textsuperscript{m} variant produced only the unspliced transcript. For both the SD\textsuperscript{wt} and the SD\textsuperscript{m} virus, an increased level of these RNAs was detected with a larger Tat input, which is in agreement with activation of transcription by Tat.
Increasing the Tat concentration also resulted in a higher CA-p24 level for both viruses. Notably, co-transfection of 1000 ng SD$^{\text{wt}}$ with 5 ng pTat resulted in a 4-fold increased CA-p24 level, while a 15-fold increase was observed for the SD$^{\text{m}}$ variant. A similar effect was seen with 50 ng pTat, which resulted in a 5 and 161-fold increase in CA-p24 production of the SD$^{\text{wt}}$ and SD$^{\text{m}}$ viruses, respectively. To exclude that these differences in fold-increase are caused by the much higher basal CA-p24 production of the SD$^{\text{wt}}$ virus (in the absence of Tat), we also tested the Tat effect when the cells were transfected with smaller amounts of the wt construct (10 to 333 ng SD$^{\text{wt}}$; Fig. 6B). Reducing the input of the SD$^{\text{wt}}$ construct significantly reduced basal and Tat-activated CA-p24 production, but did not significantly affect the fold-increase in CA-p24 production upon Tat co-transfection, as we observed a 2 to 5-fold increase at 5 ng pTat and 5 to 10-fold increase at 50 ng pTat. From this it can be concluded that the relatively low Tat induction of the SD$^{\text{wt}}$ virus is not due to its high basal expression level. More likely, this low induction is caused by the Tat splicing effect. Thus, the strong induction observed for the SD$^{\text{m}}$ virus reflects the stimulatory effect of Tat on transcription, while the lower induction level observed for the SD$^{\text{wt}}$ virus reflects the combined effects of Tat on transcription and splicing, which increases and reduces CA-p24 production, respectively. These results confirm the stimulatory effect of Tat on splicing at the major 5'ss in the complete HIV-1 context.

**Figure 6. Tat activates HIV-1 splicing.** (A) Tat and Rev-deficient HIV-1 LAI virus variants with wt or inactivated major 5'ss were constructed. A combination of one forward and two reverse primers was used in the RT-PCR analysis of the viral RNAs. This RT-PCR will result in a product of 218 bp for the spliced RNA and 667 bp for the unspliced RNA. In theory, the unspliced RNA may additionally yield a 4423 bp product, but this product will not be efficiently amplified and detected due to its large size. (B) 293T cells were transfected with 10 to 1000 ng SD$^{\text{wt}}$ or 1000 ng SD$^{\text{m}}$ HIV-1 constructs and different amounts (0 to 500 ng) Tat plasmid. The CA-p24 protein level was measured in the culture supernatant after 48 h. The mean and SD of 2 experiments is shown. (C) The total intracellular RNA was isolated after 48 h and analyzed by RT-PCR to detect the unspliced and spliced RNAs. Because the SD$^{\text{m}}$ variant produces less RNA transcripts, this part of the gel was longer exposed to allow visualization of the RT-PCR products and to demonstrate that this variant does not produce spliced RNAs.
Discussion

Transcription and splicing are tightly coupled processes (19). Several studies suggested a role for cellular transcription factors (e.g. TCERG1/CA150, Tat-SF1) in HIV-1 splicing and a role for cellular splicing factors (e.g. SKIP, SRSF1) in HIV-1 transcription (21, 36, 38, 39). We here demonstrate that the viral Tat protein, a strong activator of viral transcription, also stimulates splicing at the major 5'ss that is located in the untranslated leader RNA and that is used for the generation of all spliced RNAs. Like for the transcription mechanism, Tat requires the viral LTR promoter and the TAR hairpin for the splicing effect, which indicates that both activities are closely intertwined.

The two prevailing models for the coupling between transcription and splicing propose that either splicing factors are recruited to the promoter (recruitment model) or that the RNAP II elongation rate regulates the availability of splice sites to the splicing machinery (kinetic model) (52, 53). The C-terminal domain of RNAP II has a central role in both models, as differential phosphorylation of this domain is thought to determine both the association of transcription and RNA processing factors and to regulate RNAP II processivity. At the HIV-1 LTR promoter, this phosphorylation is mediated by the p-TEFb complex that is recruited in an inactive, 7SK snRNP bound state. Binding of Tat to the nascent TAR RNA and P-TEFb is thought to result in displacement of 7SK snRNP and activation of P-TEFb, which results in a phosphorylated and more-active RNAP II. This Tat effect on RNAP II phosphorylation can also explain the Tat effect on splicing, either by the recruitment or kinetic model of coupled transcription/splicing. In agreement with this explanation, we observed that both the HIV-1 LTR promoter and TAR element are required for the Tat splicing effect. A simplified model illustrating the combined stimulatory effects of Tat on transcription and splicing is shown in Figure 7.

Jablonski et al. previously demonstrated that Tat activates splicing to the 3’ss sites of the Rev and Env/Nef specific mRNAs (41). Surprisingly, this Tat effect required the TAR hairpin, but not the HIV-1 LTR promoter as it was also observed with non-viral promoters. The latter result suggests that this Tat splicing effect can be uncoupled from the transcription effect. Their analyses indicated the involvement of Tat-SF1 and TCERG1 (CA150) that are recruited to the RNAP II complex and stimulate assembly of the SRSF1 (SF2) splice factor onto the GAR splicing enhancer downstream of the Rev and Env/Nef specific 3’ss. This interaction promotes the upstream 3’ss sites and recruitment of U1 snRNP to the downstream 5’ss, which - by an unknown mechanism - leads to increased expression of the Env-encoding mRNA (41). Although the Tat splicing effect that we describe at the major 5’ss requires the HIV-1 LTR promoter, unlike the 3’ss results described by Jablonski et al., a similar mechanism that involves recruitment of Tat-SF1, TCERG1 (CA150) and splicing factors may be involved, as 5’ss splicing is influenced by several nearby positioned splice regulatory elements (15, 54).
Figure 7. Tat activates HIV-1 transcription and splicing. Schematic illustrating of the intertwined Tat effects on transcription and splicing. (A) In the absence of Tat, a non-processive RNAP II transcription complex is formed at the 5’LTR promoter that contains inactive, 7SK snRNP-bound P-TEFb (consisting of CDK9 and CycT1). RNAP II pausing at the stable TAR hairpin RNA structure in the nascent transcripts causes premature termination of transcription. (B) Binding of Tat to P-TEFb and TAR RNA results in the recruitment of P-TEFb to TAR and displacement of 7SK RNA, which activates the CDK9 component to phosphorylate RNAP II. The activated phosphorylated RNAP II complex is not paused by the TAR structure, which results in the production of full length primary transcripts. (C) The phosphorylated RNAP II complex may also recruit splicing factors (SF) that subsequently stimulate splicing at the major 5’ss. In this simplified model only the mentioned factors are drawn.

Both the 5’ and 3’ terminal repeat (R) region in the HIV-1 full-length primary transcript encodes the polyadenylation (pA) signal. Polyadenylation at the 5’ site has to be suppressed as this would result in short transcripts and hamper virus replication. Several mechanisms may be involved in the selective activation of the 3’ pA site (55-61), including a role for suppressive RNA structure (62) and inhibition of premature polyadenylation at the 5’ pA site by the binding of U1 snRNA to the downstream-positioned major 5’ss (46, 47). In our study, we used LTR-leader-luc and complete HIV-1 constructs with mutations in the 5’ss region (SDm) that prevent U1 snRNA binding. These mutations did indeed activate polyadenylation at the 5’ pA site in the HIV-1 transcripts but not in the LTR-leader-luc RNAs, which suggests the involvement of additional context-dependent factors. Ashe et al. (47) described that the effect of 5’ss mutation on 5’ pA site use in HIV-1 derived reporter constructs depends on the 3’ss strength. 5’ss mutation activated the 5’ pA site in the presence of a weak 3’ss, but such an effect was not observed with a strong 3’ss, which was explained by activation of cryptic 5’ss sites that took over the inhibitory effect of the regular 5’ss. However, a cryptic and likely weak 3’ss is used for splicing in the
LTR-leader-luc constructs (17). RT-PCR analysis of the RNAs produced by LTR-leader-luc SD<sup>m</sup> did not indicate activation of any cryptic 5’ss, as no differently sized RNA products were detected (Fig. 3B). Another context specific factor, e.g. a strong SV40-derived pA site at the 3’ end of the LTR-leader-luc transcripts that can efficiently compete with the 5’ pA site in the leader, may be responsible for the observation that SD<sup>m</sup> mutation did not activate 5’ pA usage, but this was not investigated further.

Tat is an attractive target for anti-HIV-1 drug therapy because it not only activates transcription but also splicing and modulation of either activity will inhibit virus replication (18, 63-65). Both activities require the specific binding of Tat to TAR and several drug candidates have been developed that prevent this interaction (66-68), including substituted guanidine indole derivatives and artificial nucleobase-amino acid conjugates (69, 70). Better understanding of how Tat regulates usage of 5’ and 3’ss sites may identify novel targets for antiretroviral drugs.

Acknowledgements
This research was sponsored by the Netherlands Organisation for Scientific Research (Chemical Sciences Division; NWO-CW; TOP grant). We thank Alex Harwig for critical reading of the manuscript.

Material and Methods
DNA constructs
The LTR-leader-luc plasmid (pLTR-gag-flag-luc<sup>6wt</sup>) and derivatives with a mutated polyadenylation site (pA<sup>m</sup>) or a mutated SD site (SD<sup>m</sup>; a.k.a. J1) were described previously (14, 17, 18). To combine the pA<sup>m</sup> and SD<sup>m</sup> mutations, the HindIII-NcoI fragment from pLTR-gag-flag-luc-SD<sup>J1</sup> that contained the SD<sup>m</sup> mutation was ligated into the corresponding sites of pLTR-gag-flag-luc-pA<sup>m</sup>. To construct an LTR-2ΔtetO-leader-luc variant with an inactivated TAR element (TAR<sup>m</sup>) and tetO elements in the U3 promoter region, we inserted the HindIII-NcoI leader fragment from pLTR-gag-flag-luc into the corresponding sites of plasmid pLTR-2ΔtetO-luc (71) that contains the HIV-1 LTR promoter with 2 tetO elements between the NFκB and Sp1 sites, an inactivated TAR element and a truncated leader region.

In the plasmid HIV-LTR-7dtetO-n-CMV2-dNF-luc (to be published elsewhere), 7 tetO sequence elements and the 30-bp minimal TATA-box containing promoter region derived from the cytomegalovirus immediate early (CMV IE) promoter replaced the HIV-1 U3 promoter region in pLTR-2ΔtetO-luc. To introduce the wt and SD-mutated (SD<sup>m</sup>) HIV-1 RNA leader region in this construct, the leader region in the pLTR-gag-flag-luc and pLTR-gag-flag-luc-SD<sup>J1</sup> plasmids was PCR amplified with the primers CMV-TATA-mut2 (<i>GAAAGTCGACTATATAAGCAGAGCTCGTTTAGTGAACCAGG</i>TCTCTCTGGTTAGACG; AgeI site underlined; TAR sequence in italics) and the reverse luciferase primer TA016 (14). The resulting PCR fragments were digested with HindIII and Ncol and ligated into the corresponding sites of HIV-LTR-7dtetO-n-CMV2-dNF-luc, which resulted in the tetO-leader-TAR<sup>m</sup>/SD<sup>wt</sup>-luc and tetO-leader-TAR<sup>m</sup>/SD<sup>m</sup>-luc constructs.

The Tat and Rev open reading frames in the HIV-1 pLAI molecular clone were inactivated by mutagenesis PCR, using the overlap extension method (72) and pLAI-Tat<sup>stop</sup> (Harwig et al. submitted for publication) as template. PCR fragment 1 was produced with primers Vpr-primer1-Not1 (<i>ACCGCGGCGCAAGGGCCACAGGGA</i>) and Tat-splice1 (51). PCR fragment 2 was produced with
primers Rev-Augstop1 (AGGCATCTCCTTATGGCAGGAAGACTGAGACAGCGAGACAGCGA) and WS3 (51). The fragments 1 and 2 were combined and used as template in a third PCR with the outer primers Vpr-primer1-Not1 and WS3. The resulting PCR fragment was digested with SalI and KpnI and ligated into the corresponding sites of pUHD6100-8911, a shuttle vector containing the Ncol-BamHI Tat-Vpu-Env fragment from pLAI (51). After digestion with Ncol and BamHI, the mutated Tat-Vpu-Env fragment was ligated into the corresponding sites of pLAI, which resulted in HIV-Tat"^m"-Rev"^m". To inactivate the SD site, the AgeI-BssHII leader fragment of the SD-mutated LAI-SD construct (17, 18) was ligated into the corresponding sites of HIV-Tat"^m"-Rev"^m".

The pRL-CMV plasmid contains the Renilla luciferase reporter gene under the control of the CMV-IE promoter (Promega). The plasmid pTat-exon expresses the HIV-1 LAI Tat protein under control of the CMV-IE promoter (51).

**Cell culture and transfection**

293T cells were cultured at 37°C and 5% CO₂ in DMEM medium containing 10% fetal bovine serum (FBS), nonessential amino acids, 40 units/ml penicillin and 40 µg/ml streptomycin. Cells were transfected by calcium phosphate precipitation as previously described (71).

**Luciferase activity assays**

293T cells (~1.5 x 10⁵ cells per 2 cm², 60% confluency) were transfected with 20 ng of the LTR-leader-luc or tetO-leader-luc construct, 0.5 ng pRL-CMV (as internal control), 5 ng pTat-exon (51) or the empty pcDNA3 vector, and 0.4 ng pCMV-rtTA-V1 (71) when indicated. The DNA was completed to 500 ng with pBluescript as carrier DNA. When indicated, 1000 ng/ml dox was added to the culture medium to activate rtTA. The cells were cultured for 48 h and subsequently lysed in passive lysis buffer (Promega). Firefly and Renilla luciferase activities were determined with the dual-luciferase assay (Promega) and the firefly luciferase activity was normalized to the Renilla luciferase activity. Data were corrected for between-session variation with the factor correction software (73).

**RT-PCR analysis**

293T cells (~1.5 x 10⁵ cells per 2 cm², 60% confluency) were transfected with 1 µg of the LTR-leader-luc or tetO-leader-luc construct, 50 ng pTat-exon or pcDNA3, and 0.4 ng pCMV-rtTA-V1 (71) when indicated. The cells were cultured for 48 h (with 1000 ng/ml dox when indicated) and washed with phosphate-buffered saline (PBS). Total cellular RNA was isolated with the RNeasy kit (QIAGEN) and contaminating DNA was removed with RNase-free DNase during isolation (QIAGEN). The RNA was used as template for cDNA synthesis with oligo-dT and random hexamers as primer (ThermoScript RT-PCR system; Invitrogen). The cDNA product obtained with the LTR-leader-luc constructs was PCR amplified with primers TA033 (74) and TA113 (17) as described (17). The cDNA product for the tetO-leader-luc constructs was PCR amplified with primers LTRalu1 (CTGGCTAACTAGGGAACCCACTG) and TA113 (17).

**Analysis of HIV-1 splicing and transcription**

293T cells were transfected with 1 µg of the HIV-Tat"^m"-Rev"^m" construct, 0 to 500 ng pTat-exon and 0 to 500 ng pcDNA3 (in total 1.5 µg DNA). The culture supernatant was harvested after 48 h and virus production was measured by CA-p24 ELISA (75). The intracellular RNA was isolated and analyzed by RT-PCR as described for the LTR-leader-luc transfected cells, using a combination of one forward (5’ triple-Gag, GACGCAGGACTGCTGGTCAGTG) and two reverse PCR primers (circ9, TTAGCATGGTGTTTAAATCT; 3’Pol1Br37, GCCACACAATCATCAGCTGCGCACTCG) to detect both the unspliced and Vif-encoding spliced transcript.

**References**


37. Miller HB, Saunders KO, Tomaras GD, Garcia-Blanco MA. 2009. Tat-SF1 is not required for Tat transactivation but does regulate the relative levels of unspliced and spliced HIV-1 RNAs. PloS one 4:e5710.


47. Ashe MP, Pearson LH, Proudfoot NJ. 1997. The HIV-1 5' LTR poly(A) site is inactivated by U1 snRNP interaction with the downstream major splice donor site. The EMBO journal 16:5752-5763.


